AL-TP-1992-0063

ARMSTRONG LABORATORY

PILOT CANDIDATE SELECTION METHOD (PCSM): WHAT MAKES IT WORK?

Thomas R. Carretta
Malcolm James Ree

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
7909 Lindbergh Drive
Brooks Air Force Base, TX 78235-5352

March 1993

Interim Technical Paper for Period January 1992 - December 1992

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

Thomas R. Carretta

THOMAS R. CARRETTA
Project Scientist

WILLIAM E. ALLEY, Ph.D.
Technical Director
Manpower and Personnel Research Division

ROGER W. ALFORD, Lt Colonel, USAF
Chief, Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1993	Interim January 1992 – December 1992

4. TITLE AND SUBT

Pilot Candidate Selection Method (PCSM): What Makes It Work?

5. FUNDING NUMBERS

PE - 62205F
PF - 7719
TA - 18
WU - 45

6. AUTHOR(S)

Thomas R. Carretta
Malcolm James Ree

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
7909 Lindbergh Drive
Brooks Air Force Base, TX 78235-5352

8. PERFORMING ORGANIZATION REPORT NUMBER

AL-TP-1992-0063

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A sample of 678 Air Force pilot training candidates were tested with a paper-and-pencil aptitude battery and computer-administered tests of psychomotor skills, information processing, and attitude toward risk. A self report of flying experience was also collected. These data were used in regression analyses to determine which variables provided the best prediction of two flying criteria, passing-failing flying training and class ranking at the end of flying training. The paper-and-pencil tests were found to be the best predictors. The measures of flying experience, psychomotor skills, and attitude toward risk incremented the prediction of the criteria. Information processing was not found to be incremental to the other variables in the prediction of the criteria.

14. SUBJECT TERMS

Pilot candidate selection
Test validation

15. NUMBER OF PAGES

22

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRA.
Unclassified	Unclassified	Unclassified	UL

CONTENTS

List of Tables

PREFACE

This research and development effort was conducted under Work Unit 77191845 which is dedicated to the selection and classification of USAF aircrew personnel. The authors thank Maj David Perry, Dr. Joseph L. Weeks, and Dr. William E. Alley for their comments.

PILOT CANDIDATE SELECTION METHOD (PCSM): WHAT MAKES IT WORK?

INTRODUCTION

Modern high-performance jet aircraft place heavy demands on Air Force pilot's physical condition, psychomotor coordination, and cognitive/perceptual abilities. The identification of candidates most likely to succeed as Air Force pilots has been a long standing goal (Bordelon & Kantor, 1986; Carretta, 1989,1990, 1992; Hunter & Thompson, 1978; Long & Varney, 1975; McGrevy & Valentine, 1974; Miller, 1947; Morales & Ree, 1992; Ree, 1976; Stoker, Hunter, Kantor, Quebe, & Siem, 1987). The variables currently considered in pilot candidate selection include medical and physical fitness, college performance, paper-and-pencil aptitude test scores (e.g., Air Force Officer Qualifying Test (AFOQT); see Skinner & Ree, 1987 for a description), and previous flying experience.

Air Training Command has initiated several programs that will significantly change the process by which Air Force pilot candidates are selected, classified, and trained. The changes are a result of policy decisions (a) to convert from a generalized undergraduate pilot training (UPT) system to a specialized undergraduate pilot training (SUPT) system, (b) to classify pilot candidates into specialized training tracks (bomber/fighter or tanker/transport) at the end of T-37 (initial jet trainer) training, and (c) to operationally implement a recently validated computer-based pilot candidate selection instrument (Basic Attributes Test (BAT): see Carretta, 1987 for a description).

The Pilot Candidate Selection Method (PCSM) is the SUPT subcomponent by which the Air Force will select pilot candidates. The goal of PCSM is to identify the best qualified pilot training applicants and to reduce attrition. The PCSM algorithm combines scores from the AFOQT and BAT with previous flying experience to predict flying training performance and ranks applicants on probable success in flying training.

Several studies have demonstrated the incremental validity of the BAT when used with AFOQT and other current pilot selection measures (Bordelon & Kantor, 1986; Carretta, 1989, 1990; Kantor & Carretta, 1988). Operational implementation of PCSM is expected to begin in 1993 following purchase of BAT systems.

The purpose of this study was to determine what makes the PCSM algorithm work; that is, what are the sources of its predictive utility? A better understanding of the relationships among the PCSM components and pilot training performance is needed to facilitate development of next generation pilot candidate selection instrumen

METHOD

Subjects

The subjects were 678 pilot trainees in the United States Air Force. They were mostly male (98%), White (90%), and all were college graduates between the ages of 23 and 27. All pilot trainees had been selected for pilot training on the basis of scores on an aptitude test (AFOQT), educational attainment, physical standards, and a desire to fly. Although all trainees had the opportunity to decline participation in the study, none did.

Measures

The AFOQT is a cognitive paper-and-pencil multiple-aptitude battery. The battery is comprised of 16 tests measuring psychometric *g* (Earles & Ree, 1991) and the common factors of verbal, quantitative, spatial, perceptual speed, and aircrew aptitude/interest (Skinner & Ree, 1987). The tests are: Verbal Analogies (VA), Arithmetic Reasoning (AR), Reading Comprehension (RC), Data Interpretation (DI), Word Knowledge (WK), Math Knowledge (MK), Mechanical Comprehension (MC), Electrical Maze (EM), Scale Reading (SR), Instrument Comprehension (IC), Block Counting (BC), Table Reading (TR), Aviation Information (AI), Rotated Blocks (RB), General Science (GS), and Hidden Figures (HF). All tests were scored with number right.

The tests are aggregated into the 5 composites of Verbal, Quantitative, Academic Aptitude, Navigator-Technical, and Pilot. These composites are used in the commissioning of officers through the Reserve Officer Training Corps (ROTC) and the Officer Training School (OTS). The composites are also used to select candidates for pilot and navigator training.

The BAT is a computer-administered battery of tests measuring psychomotor skills, information processing, and attitude toward risk which has been validated for selection of candidates for pilot training (Carretta, 1989, 1990, 199[illegible]. The BAT was administered with a special alpha-numeric keypad, a monochrome monitor, and two control (joy) sticks. A detailed description of the BAT was provided by Carretta (1987).

The first psychomotor test was a rotary pursuit task called Two-Hand Coordination, an example of Fleishman's multilimb coordination (Fleishman & Quaintance, 1984). In this test the subject used right and left hand control sticks to keep a circle on a representation of an airplane as it moved in an ellipse on the computer monitor. The score was horizontal tracking distance error (THH). Complex Coordination, an example of control precision and multilimb coordination (Fleishman & Quaintance, 1984) was the second psychomotor test. Using the right hand control stick, this compensatory tracking task required the subject to keep a 1 in. cross centered on a dotted-line cross which bisected the monitor horizontally and vertically. Simultaneously, using the left hand control stick, the subject had to keep a 1 in. vertical bar horizontally centered at the base of the monitor display. The 1 in. cross and the vertical bar were forced

away from center by a random function. The three scores for this test were horizontal tracking distance error (CCH) and vertical tracking distance error (CCV) for the 1 in. cross and tracking distance error (CCR) for the 1 in. vertical bar. The third psychomotor test, Time Sharing, was identified with Fleishman & Quaintance's (1984) psychomotor factors of reaction time and rate control. In the first 10 min, the subject was required to keep randomly moving cross hairs on an airplane target using the right hand control stick. In the next 6 min the subject had to repeat the tracking task and had to cancel digits which appeared at random intervals and positions on the monitor. Cancellation was timed and consisted of pressing the corresponding digit on the numeric keypad. Tracking task difficulty was computer adjusted. Smaller tracking errors caused the stick sensitivity to increase and larger tracking errors caused it to decrease. The score on this test was tracking difficulty during digit cancellation (TSD). Electro-mechanical versions of these psychomotor tests were administered during World War II and are reported by Thorndike and Hagen (1959).

Information processing capacity was measured by Mental Rotation and Item Recognition. The Mental Rotation measure was a variation of a spatial transformation task (Shepard & Metzler, 1971) which required the subject to make a same-different judgment about two sequentially presented letters. Letter pairs were either same or mirror images and in the same orientation or rotated in relation to each other. A correct "different judgment" is associated with letters being mirror images and independent of rotation while a correct "same judgment" is associated with the letters being not mirror images and is also independent of rotation. The score on this test was average response time adjusted for accuracy (MRT). If the responses were below 75% correct, the reaction time score was set to 2,500 ms. Item Recognition was a measure of short-term memory (Sternberg, 1966) in which the subject was presented with a group of 1 to 6 numbers which was then removed from the display. A single number was then presented and the subject had to specify whether that number was among the group presented. The score (ITT) was average response time adjusted for accuracy. Again, the 75% correct rule was applied with 2,500 ms recorded for all scores below this minimum.

The Activities Interest Inventory was administered as a measure of attitude toward risk taking (Mullins, 1962) and consisted of 81 pairs of activities. Each pair contained one low-risk and one high-risk activity. The subjects chose between them and the scores were the percent of high-risk activities chosen (AIP) and the average response time (AIT) for making the choices.

A self report of the number of flying hours (FLYEX) accrued before entrance into the Air Force was collected. The criteria were pass-fail (P/F) in UPT and class ranking based on flying and academic grades (RANK) during training.

Procedures

The subjects took the BAT while attending a basic course in airmanship including flying a single engine, propeller-driven, high-wing light aircraft. They then entered UPT where the criteria were collected.

As these subjects were all selected on the basis of their AFOQT scores, educational attainment, interest, and flight screening performance, they were a range-restricted sample. This restriction artificially causes the correlations to be downwardly biased estimates and must be corrected. Lawley's (1943) multivariate correction for range restriction was applied to the matrix of correlations from the sample to make it represent the expected correlations in a group of 3,000 applicants (Skinner & Ree, 1987). As the Skinner and Ree sample did not contain correlations involving education, it is likely that the corrected matrix is still an underestimate (Linn, Harnisch, & Dunbar, 1981) of the population values. The Lawley correction could not be applied to a matrix that included both the Pilot composite and the AFOQT tests due to linear dependency among these variables. Nor could the Lawley correction be applied to the Pilot composite alone, and a series of univariate corrections (Thorndike, 1949) would be inappropriate. Therefore, the matrix of correlations including the Pilot composite and the other variables is downwardly biased and underestimates the true values of the correlations. Test scores rather than composites were used in certain analyses to afford maximum prediction.

Descriptive statistics, correlations, and regressions were computed for the sample. Correlations used to compute the regressions involving error and response time scores were reflected so that good performances were always positively correlated. To determine the predictive efficiency of types of variables, linear model analyses were conducted (Ward & Jennings, 1973). The criteria were regressed on each aggregation of variables of a specific type (i.e., AFOQT, psychomotor, information processing, attitude toward risk, and flying experience). Using pairs of full and restricted models, the incremental validity of each variable type was tested against the baseline of the operational multiple aptitude test, the AFOQT. Additionally, a regression model that contained all the variables was tested against 5 other models that contained all the variables except one type. For example, a regression equation that contained all the variables was tested against a regression equation that contained all the variables except the psychomotor variables. This test allowed for an estimate of the unique contribution of each type of variable.

RESULTS

Examination of the means and variances of the AFOQT scores showed that the sample was range restricted. The means were all higher and the variances reduced when compared to the applicant sample (Skinner & Ree, 1987). On average, the test means were increased by .59 standard deviation units. For 14 of 16 tests, the variances decreased to an average of 7(% of the variance of the applicant sample. The IC and AI tests showed an average ***increase*** in variance to 105% of the applicant sample variance. While this increase was unusual, it was found elsewhere in the literature (see Levin, 1972) and is a consequence of selection procedures. The test that showed the greatest reduction in variance was TR which is simultaneously on the Pilot and Navigator-Technical composites, both of which are used directly in pilot selection. The least variance restricted tests (not including the 2 which showed increases in variance) were DI and GS, both on the Navigator-Technical composite. These tests showed 84% of the applicant sample variance.

Due to the size of the correlation matrix, 676 entries, it is not reproduced here but is available on request. The uncorrected correlations range from low to moderate with unexpected negative correlations on the aptitude tests, due to range restriction. The corrected matrix shows less downwardly biased estimates and stronger correlations. Some of the previously negative correlations have been reestimated to be positive in keeping with the Lawley theorem (Birnbaum, Paulson, & Andrews, 1950; Lawley, 1943; Ree & Carretta, in press).

The results of the regression analyses are shown in Table 1. Almost all of the variable types were statistically significant predictors of the criteria.

Table 1. Regression Analyses Using AFOQT Pilot Composite (Uncorrected Correlations)

		R		ΔR	
Scores	N Scores	UPT (P/F)	Rank	UPT (P/F)	Rank
AFOQT Pilot	1	.168**	.200**		
BAT Psychomotor	5	.148*	.158*		
BAT Info Proc[a]	2	.058	.027		
BAT Risk[a]	2	.101*	.108*		
Flying Experience	1	.167**	.19[illegible]**		
Pilot and Psychomotor	6	.207**	.238**	.039	.038*
Pilot and Cognitive	3	.174**	.206**	.006	.006
Pilot and Risk	3	.203**	.236**	.035**	.036**
Pilot and Flying Experience	2	.235**	.274**	.067**	.074
All	11	.295**	.333**	.127**	.133**

[a] Info Proc is information processing and Risk is attitude toward risk.
*P < .05
**P < .01

Incremental validity of the predictors beyond the prediction offered by the AFOQT Pilot composite can be found in the last 2 columns. The predictor with the greatest incremental validity was flying experience. The type of predictor with the least incremental validity was information processing. Incremental validity of the predictors were: psychomotor, .039 for P/F and .038 for RANK, information processing, .006 for both criteria, attitude toward risk, .035 and .036 for P/F and RANK, and flying experience showed the greatest incremental validity at .067 and .074 for P/F and RANK. The incremental validity of all the variables beyond the Pilot composite was .127 and .133 for P/F and RANK as criteria. The same regressions were computed using the 16 AFOQT tests, and the results are presented in Table 2. The results of the linear models analyses where one type of variable was removed and

compared to all the remaining variables are presented in Tables 3 and 4. The results presented in Tables 2, 3, and 4 closely parallel the results presented in Table 1.

Table 2. Regression Analyses Using AFOQT Tests (Uncorrected Correlations)

Scores	N Scores	R		ΔR	
		UPT (P/F)	Rank	UPT (P/F)	Rank
AFOQT Tests	16	.244**	.277**		
BAT Psychomotor	5	.148**	.158*		
BAT Info Proc[a]	2	.058	.027		
BAT Risk[a]	2	.101*	.108*		
Flying Experience	1	.167**	.190**		
AFOQT and Psychomotor	21	.268**	.302**	.024	.025
AFOQT and Info Proc	18	.247**	.280**	.003	.003
AFOQT and Risk	18	.268**	.307**	.024*	.030**
AFOQT and Flying Experience	19	.291**	.330**	.047**	.053**
All	26	.332**	.375**	.088**	.098

[a]Info Proc is information processing and Risk is attitude toward risk.
*P < .05
**P < .01

Table 3. Uniqueness Analyses Using AFOQT Pilot Composite (Uncorrected Correlations)

Scores	N Scores	R		ΔR	
		UPT (P/F)	Rank	UPT (P/F)	Rank
1. All	11	.295**	.333**		
2. All, except Pilot	10	.261**	.282**	.034**	.051**
3. All, except Psychomotor	6	.251**	.287**	.044**	.046**
4. All, except Info Proc[a]	9	.292**	.332**	.003	.001
5. All, except Risk[a]	9	.283**	.321**	.012	.012
6. All, except Flying Experience	10	.244**	.277	.051**	.056**

[a]Info Proc is information processing and Risk is attitude toward risk.
*P < .05
**P < .01

Table 4. Uniqueness Analyses Using AFOQT Tests (Uncorrected Correlations)

Scores	N Scores	R		Δ R	
		UPT (P/F)	Rank	UPT (P/F)	Rank
1. All	26	.332**	.375**		
2. All, except AFOQT Tests	10	.261**	.28[illegible]**	.071**	.093**
3. All, except Psychomotor	21	.301**	.342**	.031*	.033**
4. All, except Info Proc[a]	24	.331**	.375**	.001	.000
5. All, except Risk[a]	24	.323**	.364**	.009	.011
6. All, except Flying Experience	25	.296**	.335**	.036**	.040**

[a]Info Proc is information processing and Risk is attitude toward risk.
*P < .05
** P < .01

Regressions were also computed from the matrix of corrected correlations using the AFOQT tests and the other variables. Ree, Earles, & Teachout (1992) have shown that although the standard error of corrected correlations is not precisely known, the significance test associated with the difference between linear models is unaffected by the Lawley correction. The F test associated with the difference between linear models uses only error sums of squares which are not changed by the correction.

Table 5 shows the regressions from the corrected matrix of correlations. The corrected multiple regressions of the P/F and RANK criteria on the AFOQT tests were .308 and .347, respectively. Flying experience added the largest increment to the tests at .036, for P/F, and .041, for RANK. Increments of .019 (P/F) and .023 (RANK) were found for the measures of attitude toward risk in the corrected matrix. Adding the psychomotor scores from Two-Hand Coordination, Complex Coordination, and Time Sharing, incremented the validity of the AFOQT tests .018 and .019 for the two criteria P/F and RANK. The incremental validity of the information processing tests was .00[illegible] for both criteria. The increments above the AFOQT tests provided by using all the variables was .071 for P/F and .079 for RANK.

Table 5. Regression Analyses Using AFOQT Tests (Corrected Correlations)

Scores	N Scores	R UPT (P/F)	R Rank	ΔR UPT (P/F)	ΔR Rank
AFOQT Tests	16	.308**	.347**		
BAT Psychomotor	5	.182**	.192*		
BAT Info Proc[a]	2	.103*	.084		
BAT Risk[a]	2	.093*	.099*		
Flying Experience	1	.166**	.187**		
AFOQT and Psychomotor	21	.326**	.366**	.018	.019
AFOQT and Info Proc	18	.310**	.349**	.002	.002
AFOQT and Risk	18	.327**	.370**	.019*	.023**
AFOQT and Flying Experience	17	.344**	.388**	.036**	.041**
All	26	.379**	.426**	.071**	.079**

[a]Info Proc is information processing and Risk is attitude toward risk.
*P < .05
**P < .01

It is appropriate to remember that these regressions and increments are susceptible to shrinkage on cross application and we have calculated the expected cross validity by application of Stein's operator (Kennedy, 1988). The expected cross validity of the corrected correlations decreased by no more than .002, a trivial amount.

The results of removing one variable type and testing its uniqueness for prediction of the criteria were consistent with the linear models analyses. Tables 3, 4, and 6 show these results.

Removing flying experience from the regression containing all the variables (using the Pilot composite; see Table 3) caused the largest drops in predictive efficiency, .051 (P/F) and .056 (RANK). In both the uncorrected (Table 4) and corrected (Table 6) matrices, removal of the AFOQT tests caused the largest decrements.

Table 6. Uniqueness Analyses Using Corrected Correlations

Scores	N Scores	R		ΔR	
		UPT (P/F)	Rank	UPT (P/F)	Rank
1. All	26	.379**	.426**		
2. All, except AFOQT Tests	10	.288**	.307**	.091**	.119**
3. All, except Psychomotor	21	.353**	.399**	.026*	.027**
4. All, except Info Proc[a]	24	.378**	.426**	.001	.000
5. All, except Risk[a]	24	.371**	.416**	.008	.010
6. All, except Flying Experience	25	.349**	.392**	.030**	.034**

[a]Info Proc is information processing and Risk is attitude toward risk.
*P < .05
**P < .01

DISCUSSION

Although the information processing tests were not incremental to either the AFOQT Pilot composite or AFOQT tests or the other variables, they have been found to be incremental in a previous sample (Carretta, 1992). The reason for their lack of incremental validity may be the rather severe disproportionality (85.7% passed flying training) of the P/F criterion in this sample which is a subset of the sample in which they were previously found to be incremental. The difference between the two samples aside from the split proportions was the requirement that the current sample contain the RANK criterion for each subject. Under circumstances of less criterion disproportionality, they seem to be incrementally valid predictors.

The relatively low incremental validity of the psychomotor tests is consistent with previous findings (Ree & Carretta, in press) which showed them to be *g*-loaded. They did, however, offer unique predictive efficiency not provided by other variables.

That flying experience was the most incrementally predictive variable came as no surprise (Stoker, Hunter, Kantor, Quebe, & Siem, 1987). Additionally, removing flying experience from the models with all the variables (Pilot composite used) lead to the greatest decrement in predictive efficiency. Flying training exposes individuals to information about aircraft and may serve as a screening device to weed out those with the least motivation, those who engender fear of flying, and those who cannot learn to handle the aircraft properly. However, flying training is expensive and may also screen out potentially successful pilots due to lack of income or opportunity to pursue flying training.

Attitude toward risk (AIP, AIT) was incrementally valid beyond both the AFOQT Pilot composite and the 16 AFOQT tests. However, what it truly measures cannot be said, but its incremental validity compels further study. This test should be administered as part of a factor reference study among a series of personality marker tests. Further, its susceptibility to faking and providing responses which are socially desirable should be evaluated.

The greatest loss in prediction was found when the 16 tests of the AFOQT were removed from regressions containing all variables. These regression models were not without their problems, though. Operationally the Air Force uses the Pilot composite although other options could be considered. Many of the regression coefficients were negative and in application this would cause problems. Some of the variables would be easy to compromise by not responding to them. Also, some of the negative weights would penalize the good performance encouraged by the test administration instructions.

The paper-and-pencil tests were the most predictive variables. Flying experience, psychomotor, and attitude toward risk all contributed to the prediction of the criteria. Information processing failed to be a valid predictor and should be evaluated for revision or discarded.

REFERENCES

Birnbaum, Z. W., Paulson, E., & Andrews, F. C. (1950). On the effects selection performed on some coordinates of a multi-dimensional population. *Psychometrika, 15*, 191-204.

Bordelon, V. P., & Kantor, J. E. (1986). *Utilization of psychomotor screening for USAF pilot candidates: Independent and integrated selection methodologies* (AFHRL-TR-86-4, AD-A170 353). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Carretta, T. R. (1987). *Basic Attributes Test (BAT) system: The development of an automated test battery for pilot selection* (AFHRL-TR-87-9, AD-A185 649). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Carretta, T. R. (1989). USAF pilot selection and classification systems. *Aviation Space and Environmental Medicine, 60*, 46-49.

Carretta, T. R. (1990). Cross-validation of experimental USAF pilot training performance models. *Military Psychology, 2*, 257-264.

Carretta, T. R. (1992). Recent developments in U. S. Air Force pilot candidate selection and classification. *Aviation Space and Environmental Medicine, 63*, 1112-1114.

Earles, J. A., & Ree, M. J. (1991). *Air Force Officers Qualifying Test Estimating th general ability component* (AFHRL-TP-1991-0039, AD-A245 078). Brooks AFB TX. Manpower and Personnel Research Division, Human Resources Directorate, Armstrong Laboratory.

Fleishman, E. A., & Quaintance, M. K. (1984). *Taxonomies of Human Performance: The description of Human Tasks.* Orlando, FL: Academic Press.

Hunter, D. R., & Thompson, N. A. (1978). *Pilot selection system development* (AFHRL-TR-78-33, AD-A058 418). Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Kantor, J. E., & Carretta, T. R. (1988). Aircrew selection systems. *Aviation Space and Environmental Medicine, 59,* A32-A38.

Kennedy, E. (1988). Est ation of the squared cross-validity coefficients in the context of best subtest gression. *Applied Psychological Measurement, 12,* 231-237.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulas. *Proceedings of the Royal Society of Edinburgh.* Section A, 62 Part I, 28-30.

Levin, J. (1972). The occurrence of an increase in correlation by restriction of range. *Psychometrika, 37,* 93-97.

Linn, R. L., Harnisch, D. L., & Dunbar, S. B. (1981). Correcting for range restriction: An empirical investigation of conditions resulting in conservative corrections. *Journal of Applied Psychology, 66,* 655-663.

Long, G. E., & Varney, N. C. (1975). *Automated pilot aptitude measurement system* (AFHRL-TR-75-58, AD-A018 151). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

McGrevy, D. F., & Valentine, L. D., Jr. (1974). *Validation of two aircrew psychomotor tests* (AFHRL-TR-74-4, AD 777 830). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Miller, N. E. (1947). *Psychological Research on Pilot Traini* Report No. 8, Army Air Forces Aviation Psychology Program Research Reports. Washington D. C.: US Government Printing Office.

Morales, M., & Ree, M. J. (1992). *Intelligence predicts academic and work sample training performance.* A paper presented at the annual meeting of the American Psychological Society, San Diego, CA.

Mullins, C. J. (1962). *Objective tests of self-confidence* PRL-TM-62-6. Lackland AFB, TX: Selection and Classification Branch, Personnel Research Laboratory.

Ree, M. J. (1976). *The effects of item-option weights on the reliability and validity of the AFOQT as used for pilot selection* (AFHRL-TR-76-76, AD-A035 732), Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Ree, M. J., & Carretta, T. R. (in press). The correlation of cognitive and psychomotor tests.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1992). *General cognitive ability predicts job performance* (AL-TP-1991-0057, AD-A245 099). Brooks AFB, TX: Manpower and Personnel Research Division, Human Resources Directorate, Armstrong Laboratory.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and Factor Analysis* (AFHRL-TR-86-68, AD-A184 975). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Sternberg, S. (1966). High speed scanning in human memory. *Science, 153*, 652-654.

Stoker, P., Hunter, D. R., Kantor, J. E., Quebe, J. C., & Siem, F. M. (1987). *Flight screening program effects on attrition in undergraduate pilot training* (AFHRL-TP-86-59. AD-A183 446). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Thorndike, R. L. (1949). *Personnel Selection.* New York: Wiley.

Thorndike, R. L., & Hagen, E. (1959). *Ten thousand careers.* New York: Wiley.

Ward, J. H., & Jennings, E. (1973). *Introduction to linear models*. Englewood Cliffs, NJ: Prentice-Hall.